DARA: Assisting Drivers to Reflect on How They Hold the Steering Wheel

Dimitrios Raptis, Joakim Iversen, Tobias H. Molbak, Mikael B. Skov
Department of Computer Science / Human-Centred Computing Group
Aalborg University, Aalborg, Denmark
raptis@cs.aau.dk, joakim@joakimi.dk, tobias@molbak.org, dubois@cs.aau.dk

ABSTRACT
This paper presents DARA, the Driving Awareness and Reflection Assistant that makes drivers aware of potentially dangerous practices on how they hold the steering wheel, and helps them reflect. DARA utilizes a hand recognition component and a feedback one. The first recognizes how drivers hold the steering wheel and classifies their actions through a Leap Motion controller and machine learning. The second is comprised by a mobile application that provides drivers with feedback during and after their drive. DARA was evaluated with three studies for its accuracy, relevance and utility. Our findings show that DARA was successful both in making holding patterns present-at-hand for the drivers and in assisting them to reflect. We conclude our paper with a discussion on the broader implication of our findings on in-car hand recognition and feedback systems.

Author Keywords
Driving; Hand Recognition; Machine Learning; Feedback; Awareness; Reflection.

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous;

INTRODUCTION
Over the past decades the cars we drive everyday became less demanding and easier to drive. For many people now, driving is an established practice as the involved artifacts in use are, in Heideggerian terms, ready-to-hand [8]. This situation often allows drivers to be involved to other activities that are irrelevant to driving, such as manipulating the navigation systems, or using their mobile phones. Unfortunately, this inattentiveness is currently the most frequent reason for road accidents [17].

Dealing with inattentiveness while driving has been a focus of research for a number of years and the most common approach is to design and evaluate detection and warning systems. Such systems detect when drivers engage in inappropriate behaviors and/or are inattentive, and provide warnings. For example, most modern cars provide sound warnings when a driver is not wearing a seatbelt. In relation to recognizing specific activities that could be dangerous, such as using a mobile phone, the most common detection approaches are through monitoring drivers’ eye gaze and face orientation using a camera [e.g. 2, 3, 7, 19], through posture detection using sensors and pressure pads installed to the driver’s seat [e.g. 21], and through hand tracking using depth cameras [e.g. 18, 31, 34, 35]. In relation to warning mechanisms, various modalities have been utilized such as visual [e.g. 13], audio [e.g. 5, 33], and vibration [e.g. 4]. The challenge though is that often such systems tend to be ignored by drivers [28].

In this paper, our aim is to go beyond warning systems, and explore if and how it is possible to help drivers realize that many of their, often unintentional, actions may lead to dangerous situations. This research area is even more important now that semi-autonomous cars are being introduced to the streets, as understanding when and how to provide feedback is very relevant. For this, we designed and developed the Driving Awareness and Reflection Assistant (DARA). DARA is system that characterizes the drivers as inattentive, or attentive based on the way they hold the steering wheel, through a machine-learning enhanced, hand recognition component. We chose to focus on this, inspired by data which shows that inattentive drivers tend to have a relaxed way of holding the steering wheel [17]. After DARA characterizes a driver as inattentive, a feedback component reveals to the drivers all their intentional, or unintentional actions that can be potentially dangerous. Thus, by making drivers aware, DARA helps them reflect on their driving practices, and hopefully improve.

Our paper is structured as follows. First, we review the related work. Then we present in detail the hand recognition and feedback components, as well as, our findings from three empirical studies in relation to DARA’s accuracy, relevance and utility. Finally, we discuss the broader implication of our findings for in-car hand recognition and feedback systems.
RELATED WORK

The National Highway Traffic Safety Administration (NHTSA) of the US Department of Transportation defines distracted driving as “any activity that diverts attention from driving, including talking or texting on your phone, eating and drinking, talking to people in your vehicle, fiddling with the stereo, entertainment or navigation system—anything that takes your attention away from the task of safe driving” [17]. Furthermore, according to Stutts et al. [25], the most common secondary activities inside a car that are distracting are using the phone, eating and drinking, grooming, reading and writing, and fumbling around with objects.

These activities are characterized as secondary, because when performed people multitask while driving, and this may impact their attention to the road, as multitasking, even for small periods of time, may contribute to decreased reaction times and less control over the vehicle, exposing the driver, passengers, and the surrounding environment to unnecessary risks [22]. According to NHTSA, in 2015 3.477 people were killed and 391.000 were injured in motor vehicle crashes involving distracted drivers [17]. Furthermore, there is a clear relation between the vast majority of those secondary activities and the way drivers hold the steering wheel; in most of them drivers will hold the steering wheel only with one hand. For example, someone will hold the steering wheel with one hand in order to answer a phone call, text or manipulate the secondary controls.

Activity Recognition in a Car

A number of warning systems have been developed over the past years in an effort to help drivers become more attentive. Two are the most utilized approaches: detecting inattention by monitoring the eyes and head orientation of the driver [e.g. 2, 3, 7, 19], and activity recognition through hand tracking [31, 34, 35]. Across all the studies that deal with eye and head orientation, tracking is performed using a camera, and then, based on the collected images, a machine learning algorithm asserts the level of driver’s attentiveness. In more detail, such systems were used to detect if drivers are looking at the road [1], or how tired they are [9, 11]. In relation to hand tracking, the authors of [31, 34, 35] managed to assert drivers’ activities using depth cameras. In most of these studies, the camera is mounted on the side of the vehicle and the activities performed by the driver are classified using machine learning.

In-Car Feedback Systems

After detecting the level of attentiveness of a driver, the logical next step is to provide feedback, usually through warnings, whenever it is appropriate. In relation to how to provide feedback, three modalities have already been utilized to many cars we use every day [24]: visual, audio, and vibration. Visual feedback relies on the drivers actively diverting their eyes from the road to where the feedback is provided, and such systems have been incorporated into cars for years [13]. Christiansen et al. [5] researched in depth visual feedback systems and showed that even though less than 3% of all driver’s glances are above 2 seconds, this is enough to prove extremely dangerous while driving. When they also provided audio feedback [5], they managed to reduce the number of long glances to less than 1%, but at the same time they noted that audio feedback required more cognitive resources to process and resulted in a decrease in the driving performance.

Cao et al. [4] demonstrated that vibrations are a promising alternative to audio feedback, since it caused less interference while driving. In regards to response times and driver’s comfort though, vibrations were inferior compared to audio. In a recent project by Wang et al. [33], they explored 3D sound cues to provide spatialized feedback to represent the state outside the vehicle during critical situations. These sound cues varied in intensity and location based on the criticality of the situation. Their results showed a significant increase in drivers’ understanding and response time. At the same time though an important aspect of all in-car feedback systems is their timeliness. Research has shown that feedback provided in static intervals not only tends to be ignored, but also is deemed annoying by drivers, even during critical situations [28].

Finally, another important aspect of feedback systems is how much trust drivers have in them. For this reason, some researchers moved away from simple warnings, and the notion of personal assistants has been explored within the automotive context [15]. Their findings are similar with ones from other relevant domains, such as human-robot interaction. In this research domain, it was demonstrated that human-like and animal-like assistants not only are favored by the users [20], but are also more effective [16].

DARA: THE HAND RECOGNITION COMPONENT

Inspired by related work, we designed and evaluated the Driving Awareness and Reflection Assistant (DARA). DARA’s aim is to make the drivers aware of intentional or unintentional actions that can distract them while driving, and to help them reflect on their driving practices, in order to become better drivers. To achieve this purpose, DARA utilizes a hand recognition component, which identifies those potentially dangerous actions, and a feedback component, which informs the drivers during and after their drive. We used hand recognition as the way to identify inattentiveness inspired by related work as most distracting activities inside a car require from a driver to hold the steering wheel with one hand, or with no hands at all.

DARA’s hand recognition component utilizes a simple scenario. A recognition system tracks the driver’s hands and through machine learning, it classifies the driver’s actions. When these actions may be distracting, the driver is characterized as inattentive. In order to identify the position of driver’s hands, DARA utilizes a Leap Motion controller [14]. The Leap Motion controller was preferred from other motion recognition sensors due to its accuracy, small size, operating distance, large detection angle, and low level of
intrusiveness. To properly capture all the hand gestures of a driver, we experimented with different placements, namely next to the gearstick, on the dashboard, on the door, and on the cabin ceiling, just above the steering wheel, pointing downwards (Figure 1). Test results showed that the most efficient and least intrusive placement for the Leap Motion controller was the last one.

In order to power up the Leap Motion controller, and collect and process its data, a cable connection with a laptop was utilized (Figure 3B). Every 0.25 seconds a data sample from the Leap Motion controller is sent to the laptop and it is filtered (Figure 3C). Thus, only the relevant, above mentioned variables are fed to a Support Vector Machine (SVM) algorithm (Figure 3D). We opted for SVM because it is a supervised learning algorithm that performs well in cases with a high number of variables and large training data [27, 30]. For the classifier, we utilized the Java version of LIBSVM [6], a freely available library (Figure 3E). In order to configure the algorithm, we used as dataset, with samples from sessions that took place inside the usability laboratory, as well as, from interactions in a stationary car. The dataset was then split into training data (75%) and testing data (25%), and we repeatedly tweaked the algorithm’s values and kernel configurations until there was no impact on its accuracy. During this process, we discovered that a linear kernel with cost set of 10 provided the best results, while the remaining values were kept close to the default ones [6] (Probability=1, Gamma=1, NU=0.5, Cost=10, SVM type=Linear, Kernel type=Linear, Cache size=20000kb, Epsilon=0.001).

The Leap Motion controller provides a constant stream of data regarding the driver’s hands (Figure 3A). The data are extracted from the captured frames by its built-in infrared camera. Each frame contains a large number of variables about the direction and position of each arm, hand, and finger of the driver (Figure 2). Two were the main challenges in order to have an efficient hand recognition component for DARA: 1) to reduce the necessary number of Leap Motion variables for classifying driver’s actions, and 2) to select an appropriate machine learning algorithm and identify the best possible values and kernel configurations.

If the confidence level of the classifier is above 70%, then the hand gestures are classified and the driver is characterized as either attentive or inattentive (Figure 3F). The way DARA makes this characterization is based on related work. According to NHTSA, the ideal position to hold a steering wheel is a two-handed symmetric positioning around 9 and 3 o’clock, because it gives the most control over the vehicle and it allows for the airbag to deploy without injuring the driver’s arms. Unfortunately, though, as observational studies show [7, 19], very few drivers apply the recommended positions, even when they are fully attentive on the road. According to Jonsson [12] the most commonly observed hand position is with one hand around 10 o’clock and one around 2 o’clock. Furthermore, most drivers believe they are really in control of their car when they have both hands on the top part of the steering wheel [29].

To properly capture all the hand gestures of a driver, we experimented with different placements, namely next to the gearstick, on the dashboard, on the door, and on the cabin ceiling, just above the steering wheel, pointing downwards (Figure 1). Test results showed that the most efficient and least intrusive placement for the Leap Motion controller was the last one.

In order to address both challenges, we created an artificial driving setup in our usability laboratory, using a learning driving game on a large 50in screen, combined with a steering wheel, pedals, a car seat and a gear shift. After repeatedly using the setup, we concluded that the needed variables for our case were: 1) hand center and direction (x, y, and z), 2) hand pitch, roll, and yaw, 3) palm normalized (x, y, and z), 4) grab and pinch strength, and 5) finger direction (x, y, z) and stabilized tip position. All these values are doubles, ranging from -1.0 to 1.0, where grab and pinch range from 0 to 1.0.

In order to power up the Leap Motion controller, and collect and process its data, a cable connection with a laptop was utilized (Figure 3B). Every 0.25 seconds a data sample from the Leap Motion controller is sent to the laptop and it is filtered (Figure 3C). Thus, only the relevant, above mentioned variables are fed to a Support Vector Machine (SVM) algorithm (Figure 3D). We opted for SVM because it is a supervised learning algorithm that performs well in cases with a high number of variables and large training data [27, 30]. For the classifier, we utilized the Java version of LIBSVM [6], a freely available library (Figure 3E). In order to configure the algorithm, we used as dataset, with samples from sessions that took place inside the usability laboratory, as well as, from interactions in a stationary car. The dataset was then split into training data (75%) and testing data (25%), and we repeatedly tweaked the algorithm’s values and kernel configurations until there was no impact on its accuracy. During this process, we discovered that a linear kernel with cost set of 10 provided the best results, while the remaining values were kept close to the default ones [6] (Probability=1, Gamma=1, NU=0.5, Cost=10, SVM type=Linear, Kernel type=Linear, Cache size=20000kb, Epsilon=0.001).

In order to address both challenges, we created an artificial driving setup in our usability laboratory, using a learning driving game on a large 50in screen, combined with a steering wheel, pedals, a car seat and a gear shift. After repeatedly using the setup, we concluded that the needed variables for our case were: 1) hand center and direction (x, y, and z), 2) hand pitch, roll, and yaw, 3) palm normalized (x, y, and z), 4) grab and pinch strength, and 5) finger direction (x, y, z) and stabilized tip position. All these values are doubles, ranging from -1.0 to 1.0, where grab and pinch range from 0 to 1.0.
STUDY 1: EVALUATING DARA’S ACCURACY
The next step in our process was to evaluate the accuracy of DARA’s hand recognition component.

Setup
Participants
13 participants (4 female, 9 male) aged from 20-82 (M=37.1, SD=19.0) participated in the accuracy study using one small size car (car A) and a medium one (car B), both with manual transmission. They were all recruited through social networks and all had a valid driving license.

Process
4 participants used both cars. 8 sessions were performed in car A, and 9 in car B. The study was designed to last for 5 minutes, and it was conducted in a parking lot, where the car remained stationary. The participants were asked to perform a specific set of hand gestures: 1) hands on the steering wheel, 2) resting the hand, 3) using secondary controls, and 4) using the gear stick.

Apparatus and Data Collection
The Leap Motion controller was placed on the cabin ceiling of the car (Figure 1) and was connected to a laptop. Furthermore, each car was equipped with a video camera pointing towards the driver’s seat, ensuring that all hand gestures were recorded.

Findings
Accuracy
The collected dataset amounted for 9707 samples for car A, and 11084 samples for the car B. These samples were then compared to the video recordings. When the algorithm reported the same hand gesture as the video, the sample was characterized as accurate, otherwise as inaccurate.

The overall accuracy for the algorithm was on average 83.74% for the left hand and 87.45% for the right hand, indicating good results. In relation to specific gestures, having both hands on the steering wheel reached an average accuracy of 84.40% for the left hand, and 93.70% for the right hand. Resting the hands achieved the lowest accuracy throughout the entire test: 56.47% for the right hand, and 74.42% for the left one. DARA was also accurate in determining if the driver was manipulating the secondary controls with an average of 77.91% for the right hand. Finally, detecting the right hand on the gear stick proved to be very accurate, with an average value of 83.95% despite the gear stick being positioned close to the secondary controls and the hand gestures being similar.

Driver Variation
If we focus on the accuracy results from individual drivers, we will observe that in all hand gestures except resting the hand, DARA performed in a similar manner. For example, for detecting both hands on the steering wheel the accuracy varied between 86.54% and 99.72% among the drivers. This was not the case for resting the arm, where the accuracy varied between 4.04% and 99.9%. The reason for these large variations in accuracy is that many drivers adopt different ways of resting their hand. For example, drivers might rest their right hand by placing it on their lap, by holding the handbrake, or by placing it on an armrest. These results indicate that there are some hand gestures that are difficult to be classified from a machine learning algorithm without a large amount of training data.

Car Variation
Another interesting finding was the fact that the interior design of a car affected DARA’s accuracy. Two are the main reasons for that: a) each car utilizes different modalities for the drivers to interact with the controls (e.g. buttons, switches, rotary controls, etc.), and b) each car requires from the drivers to place their hands on a different position to perform a specific action (e.g. the air-conditioning controls may be placed on different locations within a car).

External parameters
Finally, we identified three external parameters that had a negative impact on DARA’s accuracy. The first one is the type of clothes the driver is wearing. Heavy clothes, such as winter jackets, do tend to negatively affect hand recognition. The same is the case in relation to light, as during sunny days the infrared camera of the Leap Motion controller might often not work effectively. Finally, there were some cases where the controller detected the left hand of the passenger instead of the driver’s right hand. All three parameters were experimentally controlled in the following studies.

DARA: THE FEEDBACK COMPONENT
After finalizing the hand recognition component, we designed and developed the feedback one. The feedback component is comprised by two parts: a) the during-drive feedback, where feedback is provided while driving, and b) the after-drive feedback, which is provided at the end of a trip.

Figure 4. DARA’s during-drive feedback presented to the drivers every time the car is stopped, after driving at least 600 meters.

The purpose of during-drive feedback is to inform the drivers on how they hold the steering wheel, in order to make them aware of potentially dangerous driving practices. The during-drive feedback is presented through an Android application that runs on a mobile phone, which...
is mounted on the car’s dashboard, in a similar way as a typical GPS navigation unit (Figure 4). Figure 5 displays how the during-drive feedback component works. DARA’s hand recognition component, which runs on a laptop, characterizes the drivers as inattentive or attentive every 0.25 seconds based on the hand gestures they perform (Figure 3G, Figure 5A). This information is then sent via Bluetooth from the laptop to an Android application that runs on a mobile phone (Figure 5B).

Figure 5. DARA’s during-drive feedback component.

Every 5 seconds, a new segment is created and locally stored on the mobile phone (Figure 5C). A segment contains: a) the driver’s state for each 0.25-second sample (attentive or inattentive), and b) the start and end GPS locations, acquired from the mobile phone’s GPS (Figure 5D). For each segment a score is calculated, by dividing the number of collected samples where a driver was attentive, by the total number of samples in this segment. The score ranges from 0 to 100.

When the Android application detects that the car has stopped (Figure 5E), then an overall score is calculated by averaging all the segments’ scores since the previous stop. We decided to provide during-drive feedback in the form of an overall score only when the car is stopped inspired by related work that suggests that feedback’s success is related to timeliness [28], and that feedback should be provided at times where the driver has enough time to comprehend it. Furthermore, the during-drive feedback is provided only if the car has moved more than 600 meters since the previous stop, in order not to overexpose drivers with feedback messages.

In relation to the modalities of the presented during-drive feedback, we opted for a combination of audio and visual cues. Moreover, the audio feedback utilizes a female anthropomorphic voice, since they were deemed as more trustworthy and effective in previous studies [16, 20]. The overall score is presented to the drivers for 8 seconds, along with a message (both in audio and visual form) which praises, or warns them based on their attentiveness:

- Score >= 80: “Good job on that last section”,
- 60 <= Score < 80: “Remember to hold the steering wheel properly”,
- Score < 60: “Your driving could be improved”.

When the car reaches its final destination, the after-drive feedback is provided through another view in the same Android application. Its purpose is to help drivers reflect on how they hold the steering wheel by reminding them the actions they performed during a trip.

Figure 6. DARA’s after-drive feedback provided at the end of a trip.

The after-drive feedback takes a form of a map that displays information about the whole trip (Figure 6) and was created using the Google Maps API. Each 5-second segment that is locally stored in the mobile phone contains a start and an end GPS location (Figure 5D). For each of those segments (Figure 7A), first a score is calculated, then a color is assigned (Figure 7B), and based on the start and end GPS locations, they are placed on the map (Figure 7C). Depending on the score, each segment on the map is colorized as green, yellow, or red, giving the drivers a quick overview. If drivers click on a segment on the map, they are presented with detailed information about their hand gestures and actions. For example, someone would get a low score on a segment, if he was manipulating a lot the secondary controls of the car. A final blue marker, shown at the end of the route, summarizes the entire trip.

Figure 7. DARA’s after-drive feedback component.

STUDY 2: EVALUATING DARA’S RELEVANCE

Our next step was to evaluate DARA’s relevance in a driving context. To do that, we recruited a driving instructor, provided him with DARA, and asked for his expert input. Recruiting a driving instructor also allowed us to observe in real-world conditions if our system was intrusive for the driver, since the instructor could take control of the car, if necessary.
**Process and Setup**

During the course of two hours, the driving instructor evaluated DARA with two different students, both almost ready for their driving exam. DARA was installed and calibrated for the interior of the car before starting the drive. DARA’s during-drive feedback component was installed on a mobile phone which was mounted on the car’s dashboard. The driving instructor was in charge of the route, and the small talk from a researcher that was sitting in the back seat of the car to make sure that everything was working properly, was kept to a minimum. After the driving session, we conducted a semi-structured interview with the instructor regarding his experience with DARA and we also used for input DARAs after-drive feedback.

**Findings**

*Match between DARA and the instructor*

Overall, the perception of the instructor for the hand placements of his students was accurately reflected by DARA. When he was asked to provide a score for them, he gave the first one 100 (99 by DARA), since she performed really well by properly holding the steering wheel and by quickly making gear changes. The second student was provided a 95 (96 by DARA) as there was an incident where her hands were placed very low on the steering wheel.

*Appropriateness and relevance*

Regarding the during-drive feedback, the instructor praised the timeliness and precision of the message. He felt that providing feedback only when the car was stopped, allowed the drivers to receive it and process it without being distracted by it. Concerning how we characterized a driver as inattentive, he stated that DARA nicely reflected reality, but he also suggested that a stricter hands-on-the-wheel policy should also be tested. According to him, even a driver with both hands on the top part of the steering wheel should be characterized as inattentive if he does not hold it on a 9 and 3 o’clock position.

In relation to the after-drive feedback, he found it relevant for him, his students, as well as drivers in general. He described it as a learning tool, which would allow him to have richer discussions with his students about their mistakes. Additionally, he believed that DARA could play the role of ‘virtual driving instructor’, especially for inexperienced, out-of-driving-school drivers. Finally, he also expressed his concerns that in real world conditions many drivers would probably turn DARA off, as he believed that after some time they would perceive the system as irrelevant and even annoying.

**STUDY 3: EVALUATING DARA’S UTILITY**

The last step in our process, was to conduct a field study, in order to collect data from real drivers in regards to DARA’s utility. For this study, DARA was used as a technology probe [10] in order to allow us to explore the driving context and better understand its impact.

**Participants**

10 drivers participated in the field study, and all were recruited through social networks. Following the driving instructor’s suggestion that DARA would be more relevant for inexperienced drivers, we opted mostly for them. The pool of participants consisted of 8 males and 2 females, with a mean age of 24.1 years old, and 5.45 years of actual driving experience on average. Furthermore, we asked the drivers to rate themselves on how good they believed they were driving with a single item that ranged from 1 to 10 on a Likert scale. On average, the drivers rated themselves as 7.1. Table 1 summarizes their demographic characteristics.

<table>
<thead>
<tr>
<th></th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual driving experience (in years)</td>
<td>5.45</td>
<td>1.38</td>
</tr>
<tr>
<td>Perceived driving experience (1-10)</td>
<td>7.1</td>
<td>1.10</td>
</tr>
</tbody>
</table>

**Table 1. Demographic characteristics of the participants.**

**Setup**

Each participant was asked to drive a 20km route, resulting in approximately 30min of driving. The route was comprised both by city and rural driving parts. The reason for selecting such a route was based on related work. According to statistics [32], accidents occur 15 times more often on city and rural roads than on highways. The main reason for this is that such roads are often perceived as easy by drivers, thus they divert some of their attention away from the road. Furthermore, rural driving is perceived by most drivers as calmer than city driving [32]. Thus, our participants first drove the rural part of the route in order to familiarize themselves with the vehicle as well as DARA, and then they drove the city part of the route.

From the 10 participants, half drove their own car for the field experiment, and the rest were provided a car. All cars had manual transmission. For each car, DARA was calibrated in order to accurately reflect the position of the steering wheel, the secondary controls and the gear shift. The same apparatus that was used in the previous two studies was also used in this one. The Leap Motion controller was placed on the cabin ceiling (Figure 1) and was connected to a laptop. For the during-drive feedback, a mobile phone was placed on the car’s dashboard. The same device was also used for the after-drive feedback.

**Procedure**

A test leader was responsible for introducing the participants to the route, for briefly explaining to them DARA’s purpose and the existence of a score that is based on how they hold the steering wheel. The exact mechanism on how the score was calculated was left ambiguous on purpose. The test leader also instructed the participants to drive as they would normally do and to obey the traffic laws. The test leader sat in the passenger seat next to the driver and intervened only when asked, mostly for clarifications in relation to the route.
Data Collection and Analysis
Throughout the drive, the test leader observed and kept notes on how each participant reacted to the provided during-drive feedback. These notes, along with the after-drive feedback, were used as the basis for interviews. The first part of the interview revolved around issues regarding the participants’ actions and their experience with DARA in general. During the second part, the participants were shown the after-drive feedback which showed their route, along with information about their performance. The discussion was then focused to the segments where participants did not perform well, and they were asked to recall their actions and reflect upon them.

Results
Throughout the field study, the participants’ average score ranged from 77 to 96 (M=85.4, SD=6.6). Since the during-drive feedback was provided only when the car was still, it varied in its frequency. Therefore, the participants received feedback between 5 and 11 times (M=7.4, SD=2.2). When exploring the most common gestures performed by the drivers, our data showed that they manipulated the secondary controls on average 12.89% of the time (SD=6.73), they held the gear stick on average 8.27% of the time (SD=9.33), while they rested their hands on average 5.02% of the time (SD=4.81). What was also interesting were the instances where they did not hold the steering wheel at all (M=46.2, SD=27.82 times, or M=11.55, SD=6.95 seconds).

The next step of our analysis was to compare the after-drive feedback maps of our 10 participants in order to observe if there were any road segments where they all demonstrated similar steering wheel holding patterns. Due to large variability in traffic conditions, we did not identify any such patterns, except a specific small part of the route where the road was made out of bumpy limestone blocks and they all properly held the steering wheel.

The 5 participants who borrowed a car for the field study required a few minutes to become comfortable with it, and they subconsciously drove more carefully. This was obvious both from the data collected by DARA, since all of them had really good scores in the beginning of the route (87 on average), as well as from their comments during the interviews. For example, when P1 was asked how he felt in the beginning of the route he stated:

“Well, it’s a new car, and I drive more safely and carefully when in a new car.” [P1]

Furthermore, we observed that participants who rated themselves high in regards to their driving experience, they tended to relax more while driving. This was evident by the fact they often held the steering wheel (only with one hand, or with both hands in the lower part of the steering wheel). This was also reflected in some of their scores, as the two drivers that rated themselves highest in driving experience, got relatively low scores (78 and 82), while the least experienced participant received a 95. When informed about this, they suggested they had enough control over the vehicle despite the fact they did not use the optimal hand positions.

“No, I don’t think so [About losing control due to non-optimal hand positions], but I also guess that is because I normally drive with one hand.” [P4]

Thus, we have indications that the more familiar the participants were with the car and/or the more experienced they perceived themselves, the more relaxed their driving style usually became. This resulted in more interactions with the secondary controls, more relaxed postures (such as falling back into their seat), and more reckless hand positions on the steering wheel in general.

Experiencing DARA’s during-drive feedback
Our decision not to explain in detail to the participants how their score was calculated allowed them to have a more explorative approach towards DARA. Two exploration approaches were identified. In the first one, some participants tried to find ways to lower their score in order to understand which actions were characterized by DARA as inappropriate:

“I got the score of 97. After that I wanted to see if I could get a lower score to test the system.” [P6]

In the second approach, participants that received really good scores, tried to stick to the same hand positions in order to keep the score as is. For example, as P7 mentions:

“When I got the confirmation that I was doing something good, then I tried to do the same and keep that rhythm going.” [P7]

Compared to the written message, our participants believed that the score gave a more precise measure of their performance. For example, P1 felt that the score was better, as the message was the same all the time for him, because his score was always above 80.

“The way she talks.. she can tell you that you drive well but you cannot feel the difference between 95 and 90.” [P1]

Furthermore, P3 mentioned that the score, due to its wider range, provided a better understanding on how to improve as he could experiment with different hand positions. Thus, while the message felt reassuring when the driving performance was good, the more negative versions of the message lacked information on how to actually improve the performance. As clearly highlighted by P6:

“If the system is related to where I have my hands then give feedback about where to place my hands.” [P6]

While all participants liked the fact that feedback did not distract them, as it appeared only when the car was stopped, there were different approaches in relation to its frequency.
P1, P2 and P3 asked for more feedback, and highlighted that particularly during long trips without any stops (for example in highway driving), DARA’s during-drive feedback would basically not appear at all:

“If you are just driving on a highway I don’t think you will get the message that often, unless something is incorrect or very incorrect.” [P3]

P8 moved a step further and suggested for DARA to act as a warning system and provide immediate feedback whenever the driver is not holding the steering wheel properly:

“Feedback should come as soon as a driver does something irresponsible.” [P8]

On the other hand, some participants were really satisfied with its frequency:

“Feedback came only at appropriate amount of times. It was not annoying as the GPS that spams you. It was alright.” [P9]

These findings suggest that drivers do have varying attitudes towards feedback systems and there is a need for any feedback system to adapt, to some extent, to the preferences of the driver.

Another interesting point in our findings, is that half of our participants perceived the during-drive feedback as a game. For example, P3 mentioned that he was trying to increase the score to 100 throughout the session:

“I wanted to get to 100, and see how I am supposed to drive, to learn what the best way of driving is.” [P3]

For those 5 participants, receiving a low score was challenging, and they felt that DARA encouraged them to adopt a better driving behavior, by trying to increase their score:

“Only 91! I can do better than that!” [P1]

Other gamification elements were also suggested as possible ways to extend the system, such as the possibility to save high scores and monitor the performance over a long period of time, or the possibility to compete with friends and family, or even unknown drivers by being informed about their scores.

“I could be fun to see your score compared to the rest of the drivers... How well your average drive is and how well your best drive is. Makes it more fun to get a better score.” [P10]

Such findings, are in line with related work [23, 26] that suggests that gamification principles are useful in increasing the relevance, effectiveness and engagement of in-car feedback systems. Furthermore, inspired by related work, we gave DARA an anthropomorphic voice. This decision has had an impact on how the participants experienced our system. Almost all participants stated that they felt ‘watched’ by the system while driving, and commented that this had an effect on their driving behavior. Some of them pointed out, that even though they did try to improve their score and realized the system’s value, they often had negative feelings about it. For example, P3 and P4 commented about the DARAs female voice that:

“The lady was breathing down my neck.” [P3]

“The feeling of being watched will cling to me for days.” [P4]

Finally, many participants touched upon the social implications of DARA, and especially one of them, focused a lot on the topic. On one side, he mentioned that he would feel uncomfortable to receive negative feedback in front of other people:

“It can be humiliating to get negative feedback while driving with friends.” [P3]

On the other side, he did also stress that he trusted the system way more than other passengers.

“I would rather listen to her than other passengers telling me what to do.” [P3]

These findings suggest that in-car feedback systems may bring drivers to awkward, or even embarrassing situations, and we as designers need to take into consideration the social aspects of driving.

**Experiencing DARA’s after-drive feedback**

Despite many participants driving very carefully, either due to driving a new car or due to the feeling of being watched, many of them realized through the after-drive feedback that they did unintentionally perform actions that could be dangerous. P8 and P10 felt surprised from how often they interacted with the radio. For example, P8 characterized the system as ‘creepy’ after realizing why some of his segments on the map were red, since:

“Apparently. I touched the radio more than I realized.” [P8]

Other unintentional actions that we observed during the study and led to the characterization of a driver as inattentive, were resting the right hand for too long on the gear stick, or using the hands to communicate with other passengers. When we informed P7 that this was the main reason she got a low score, she commented:

“I now am more aware of my hand positions and how much they actually, on an average drive, move.” [P7]

Through DARA, the participants realized which of their actions could lead to dangerous situations and they reflected upon them. Furthermore, when we asked them how they experienced DARA in general, most agreed that the reflections facilitated by the system could help them improve their driving practices, and confirmed that DARA
would be very relevant, especially for inexperienced drivers. For example, P5 had a personal experience of being reckless as a newly educated driver, saying that:

“At this point, I was a very irresponsible driver, but didn’t really know it.” [P8]

DISCUSSION
The main contribution of this paper is DARA, a system that helps drivers to reflect on how they hold the steering wheel by making them aware of intentional or unintentional gestures/actions that may be dangerous while driving.

We do not perceive DARA as a fully functional system that is ready to be installed in cars. As a prototype, the recognition component was accurate enough, while the during-drive feedback informed the drivers if there was something wrong with the way they held the steering wheel, and the after-drive feedback helped them realize which specific actions could be potentially dangerous. All these sparked reflections. The value that DARA may have in a driving context was also stressed by the driving instructor, who praised its relevance, particularly for inexperienced drivers. As a technology probe, DARA also allowed us to explore and better understand the driving context. In all three studies, DARA’s relevance was highlighted by how surprised many drivers were when they realized that some actions could be dangerous; when the ready-to-hand became present-at-hand [8]. For example, many were surprised on how often they interacted with the car’s secondary controls. Of course, due to the fact we conducted a small field study, we cannot make general claims about behavioral change. In order to have results on that, we need to make a longitudinal study with drivers of varying experience and then see if DARA may have a long-term impact on their driving practices.

What we deem as an important challenge for any system that helps users reflect, learn and improve a practice is how to make the involved actions present-at-hand. For this, we strongly recommend the two-step feedback approach we utilized in DARA: the during-drive and the after-drive feedback. The first one may provide an indication if there is something within a practice that needs to be improved, without enforcing to the users a specific behavior, and without interfering too much with their actions. The second type of feedback can make them aware of the actions that need to be changed, and hopefully their combination may lead to sustained changes.

Furthermore, the findings from all three studies have broader implications, which we will discuss in the following subsections.

Hand Recognition as Means to Detect In-Car Actions
Overall, DARA demonstrated that hand recognition in a car context is a viable approach. Our findings from the three studies showed that motion sensing and machine learning can be utilized in order to detect potentially dangerous actions that may decrease driver’s safety. Even though our system was designed to detect a few gestures (such as holding the gear stick, or manipulating the secondary controls), these can be easily extended. For example, driving and using a mobile phone can be very dangerous, and future implementations of DARA, or other in-car hand recognition systems, could accurately detect this, and subsequently warn the drivers.

Furthermore, our findings do also show that in order for hand recognition to be successful there is a need to calibrate the machine learning algorithms both to the internal layout of a car, as well as, (to some extend) to the individual drivers. The internal layout of the car needs to be taken into consideration as cars differ on where the secondary controls are placed. Adaptability to specific drivers may also be important since both in our accuracy study as well as the field one, we identified that some of the performed hand gestures from the drivers are too personal, while others are generic enough. For example, resting one hand while driving may occur in various ways (resting it on the gear stick, or on the door), while changing a gear is an action that is identical for most drivers.

Finally, our findings showed that DARA had trouble in properly detecting the hands of the drivers, during extremely sunny days, when the drivers were wearing heavy clothes, and sometimes when there were passengers inside the car. For our empirical studies, those three challenges were not that important as they were experimentally controlled, but in real world-settings they need to be taken into consideration.

Implications for In-Car Feedback Systems
As already has been identified in related work [28], it is important within a car context to provide timely feedback. Our decision to provide feedback only when the car was stopped received positive comments. Both the driving instructor as well as the participants in our studies did not feel distracted when feedback was presented. They felt it was presented at appropriate times and it did not affect their driving performance. The modalities used for the feedback (audio and visual) were also praised as they did not feel intrusive, or distracting.

There were though four issues that we identified as relevant for any in-car feedback system: frequency, richness, adaptability, and social dimensions. In relation to feedback’s frequency, although our participants did like the fact that feedback did not appear too often, they did express their concerns for long driving sessions where they would not receive any feedback for hours (for example while driving on a highway). For such scenarios, many participants stressed the need for the feedback mechanism to adapt its frequency. Furthermore, many participants also informed us they would prefer DARA to act as a warning system and inform them when they did something wrong, the moment they did it. This needs to be explored more though, as we know from related work that such systems as often ignored by drivers as they find them intrusive [28].
In relation to feedback’s richness, we had mixed results. Most of the participants liked the simplicity of providing just a score and a message. They felt that this information was enough for them to realize if they needed to improve the way they held the steering wheel, or if everything was fine. A few participants though stressed that they would like to have more information available during a drive. Instead of receiving only a low score, they also asked to be informed on the reasons why this happened (for example, to receive a message like ‘You spent too much time manipulating the radio’).

The third aspect of feedback that was characterized as important from most participants and the driving instructor was the need for the feedback to adapt to the driver/context. For example, it was suggested that any feedback mechanism should take into consideration how much the driver has improved over time and adapt accordingly, otherwise there is a danger to be deemed annoying and be ignored. In relation to better adapting to the driving context, we received many suggestions on how to extend DARA, by taking into consideration the traffic conditions, the weather conditions, how dangerous is the road, etc. Of course, in order to have concrete results on the effectiveness of such implementations in relation to real-time feedback systems, more studies need to be conducted, particularly in real-world settings. Nevertheless, our results do indicate that designers of similar to ours feedback systems, should take these issues into consideration.

Finally, the social dimensions surrounding feedback systems, were also highlighted by some of our participants. In a learning-how-to-drive context a feedback system like DARA can be very relevant, as it may provide to the driving instructors anchor points, which will facilitate richer discussions with the students and may improve learning outcomes. In everyday driving contexts though, receiving negative feedback in the presence of other passengers can be intrusive, embarrassing and even annoying. For this, we strongly suggest to designers of real-time feedback systems to consider such social aspects and try to facilitate alternative implementations. For example, unless an immediate danger is detected (for example in collision detection systems), during-drive feedback could be turned off in the presence of other passengers, and only after-drive feedback could be provided.

Gamification as Means to Increase Engagement
In our effort to help drivers to reflect and improve on their driving practices, we partially utilized gamification through DARA’s provided score. Half of the participants reported to us during the interviews that the score was perceived as a game, and that very often they felt they were competing. This, is in line with previous research work within driving contexts, where, for example, gamification was utilized to increase engagement and warn drivers of dangerous situations [23], or to reduce boredom and exposure to potential distractions, such as phone usage [26].

Since our minor adoption of gamification principles was more engaging for some of our participants, we would like to highlight to potential designers of feedback systems that they could introduce similar elements, and then evaluate their relevance. Examples of gamifications principles that could be considered and were discussed by our participants, was the possibility to track their progress, to compare scores with other drivers, and to utilize achievements.

CONCLUSION
In this paper, we presented DARA, the Driving Awareness and Reflection Assistant, which made drivers aware of how they hold the steering wheel as well as potentially dangerous actions they performed, and helped them reflect on them. To do so, DARA utilized a hand recognition component, and a feedback one that is comprised by during and after drive feedback.

DARA was empirically tested in three studies for its accuracy, relevance and utility. As a prototype, DARA detected accurately enough drivers’ hand placements on the steering wheel and classified them into actions. As a technology probe, it allowed us to have detailed insights on the driving context. In short, the findings from all three empirical studies, showed that the unobtrusiveness of the feedback was appreciated by drivers, since it effectively guided them to realize that many of the actions they often unintentionally perform can be dangerous. Furthermore, our findings suggest that in similar to ours in-car feedback systems, four parameters need to be taken into consideration: feedback’s frequency, richness, adaptability, and social dimensions.

As future work, we plan to extend DARA to classify more actions, and we want to study its long-term impact on driving practices, in order to understand if and how it may facilitate lasting behavioral change. Finally, we want to introduce DARA to other relevant driving contexts. First, we want to deploy it within the domain of semi-autonomous cars, where understanding how to provide timely feedback is extremely relevant. This is important, especially if we consider that there are already reports for accidents occurring due to mismatches between the drivers and the autopilots. Second, we consider extending DARA to professional drivers too, where it could be used both as a driver’s assistant, but also as a performance evaluation tool.

ACKNOWLEDGEMENTS
We would like to thank all the participants in our studies and the Trend Driving academy for their collaboration and support.
REFERENCES


5. Lars H. Christiansen, Nikolaj V. Frederiksen, Brit S. Jensen, Alex Ranch, Mikael B. Skov, and Nissanthan Thiruravichandran. 2011. Don't look at me, I'm talking to you: investigating input and output modalities for in-vehicle systems. In *Proceedings of the 13th IFIP TC 13 international conference on Human-computer interaction - Volume Part II* (INTERACT’11), 675-691. https://doi.org/10.1007/978-3-642-23771-3_49


